

本周工作：

边界项目	程序方面分为三部分，完成两部分， overview 和 specification，剩余的一部分通过模拟采样来衡量不确定性的误差对结果造成的影响。计算方法实现了，下周进行可视化 文章方面，和 ross 讨论，方向可以，根据之前提出的想法继续落实。	12 月 13
------	---	---------

边界项目：

完成了提供用户对于特定区域进行更改的操作

利用可视化界面对数据进行了展示。

实现了迭代式的 Kmeans 稳定算法而不是每次都是固定的聚类中心。更具有说服力。

遗留的问题有：Tsne 的算法不够稳定，根据每一个维度的改变完成对于每一个数值的影响的条形图，是的用户能更加清楚的了解，这个维度的数值改变多少对结果影响的大小。

和 Ross 及 yafeng 讨论项目，在不确定性的方法上面 ross 说暂时先用采样方法，其他方法下周实现后再看看。

论文：

Querying Massive Trajectories by Path on the Cloud：

对于路径的查询，以往的方法是基于单机的，不利于索引大规模的轨迹历史，轨迹更新和并发的路径查询。本文提出了一种利用 suffix tree 数据结构，在云服务器上建立路径查询系统的框架，在后端，使用预处理和索引的方法处理大规模实时数据，在前端解决大量 IO 瓶颈。使用索引处理轨迹后，在前段查询时，先把查询输入转化成索引。大大减少了查询时间。

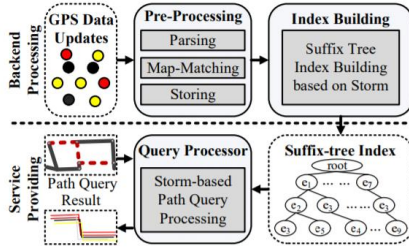


Figure 2: System Overview.

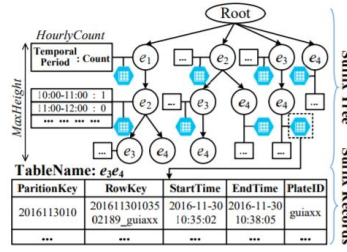


Figure 3: Table-Based Suffix Tree.

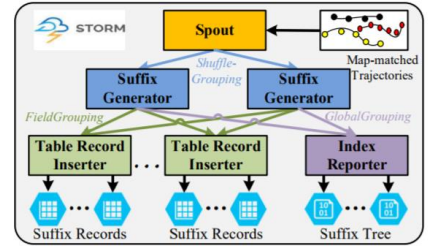
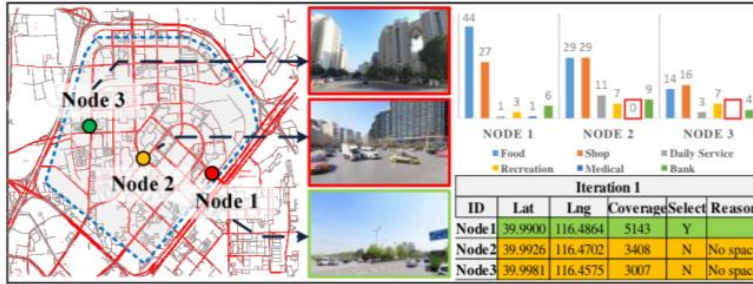


Figure 4: Storm-Based Indexing Topology.

Mining the Most Influential k-Location Set From Massive Trajectories

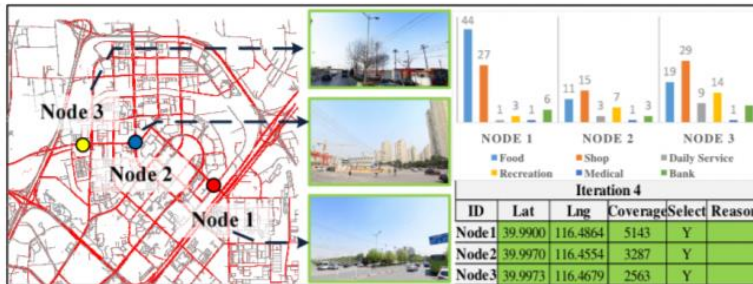
找到一个区域内最有用的地区对于寻找选点有很大帮助。本文提出了一种对于区域内选点的挖掘算法框架，对地区分块，然后计算最有影响力的区域，采用贪婪算法进行最佳优先的过滤方法，量化区域影响。例如下例子中，用户想要选择三个充电站地点，位置不能近，人要多，POI 要全。用户通过四次迭代并在每次迭代后选取保留的节点，最终完成了目标。



(a) Charging Station Suggestions in the 1st iteration

Iteration 2						Iteration 3					
ID	Lat	Lng	Coverage	Select	Reason	ID	Lat	Lng	Coverage	Select	Reason
Node1	39.9900	116.4864	5143	Y		Node1	39.9900	116.4864	5143	Y	
Node2	39.9970	116.4554	3287	Y		Node2	39.9970	116.4554	3287	Y	
Node3	39.9927	116.4704	2768	N	No space	Node3	39.9903	116.4862	2689	N	Close to 1

(b) Iterative Process in the 2nd and 3rd iterations



(c) Charging Station Suggestions in the 4th iteration

Fig. 17: Placing Charging Stations in Wangjing Area, Beijing.

Planning Bike Lanes based on Sharing-Bikes' Trajectories

本文通过启发式的贪心算法，解决 NP 难问题，寻找城市中科学的自行车道轨迹选址，方法通过城市中共享单车的轨迹数据，找到起始地 top 路段进行扩张，找到合理的自行车道建设路线。



(b) Top-k Start Segments